

German restaurants in New York

Coursera Data Science Capstone Project

Fabian Buder, 18. May 2020

Background and Motivation

Why look for German restaurants in New York

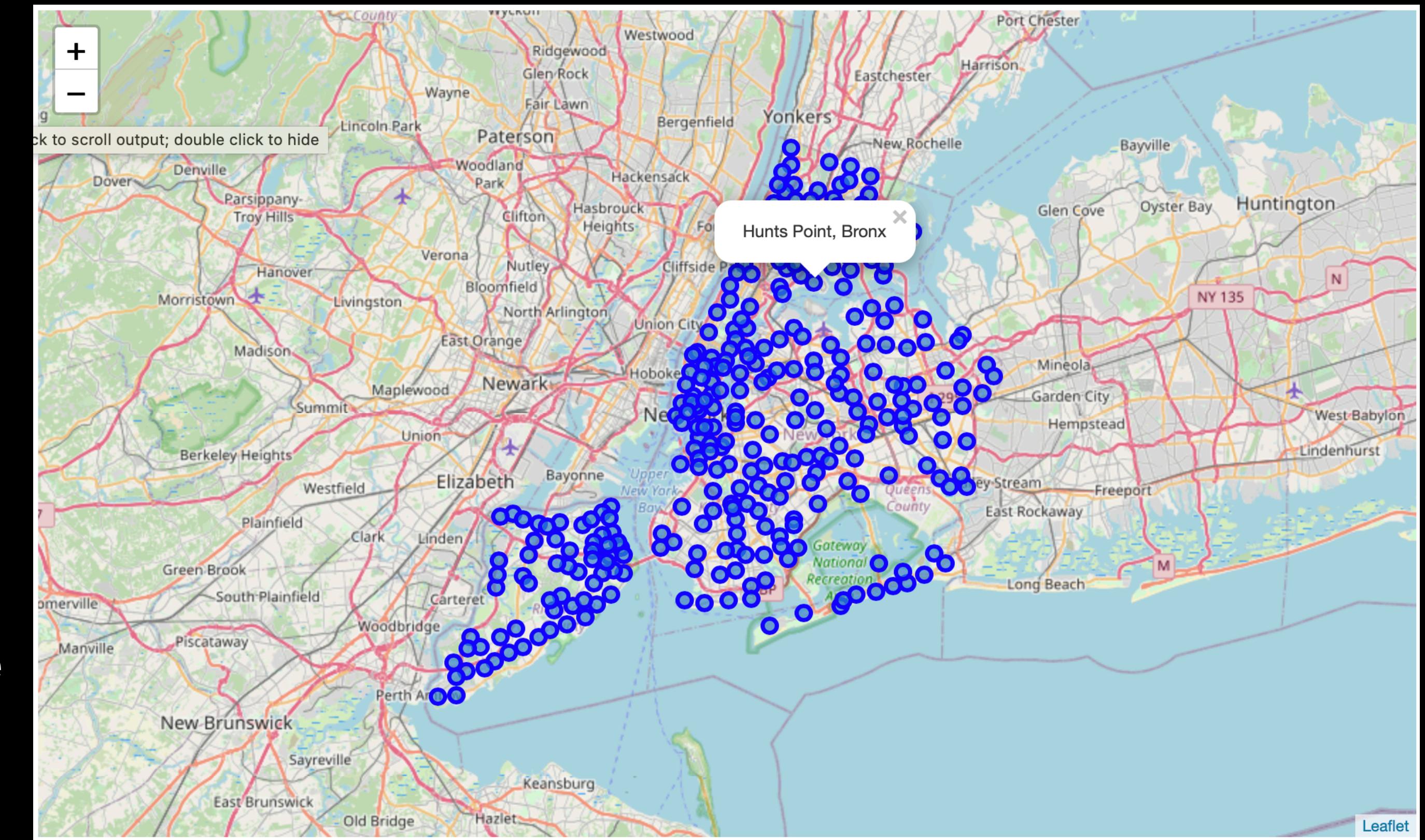
- New York is said to be one of the most exciting cities in the world. Within an area of around 300 square miles live 8.40 million people making New York also one of the most crowded places in the US. Being the dream destination for immigrants from all over the globe, New York is having an almost infinite number of restaurants serving every cuisine one can image. Even though an overall GMP of about two trillion US dollars seems like a huge potential for businesses, the market in New York is already fairly saturated and high real estate prices leave no space for the weak, at least not in the hot spots of the city where the money lies.
- For a group of investors I analyze the restaurant scene in the city "that never sleeps" to check the potential for a new German restaurant and get a first idea of what would be potential locations for such a restaurant where it would not hit too much direct competition.

Data and Methods

- I will join datasets with the geo coordinates of neighborhoods and boroughs in New York with venue data from Foursquare to analyze the competition in New York:
 - City data: https://geo.nyu.edu/catalog/nyu_2451_34572 Geo coordinates to locate the boroughs and neighborhoods in New York City.
 - Geolocation: Geopy I use the geopy package to obtain the latitude and longitude values of New York City for maps.
 - Venue information: Foursquare API The Foursquare API is utilized to obtain the locations and categories of food venues in New York.
- I will use the above mentioned data sources for maps and exploratory data analysis to give an overview over the competition and to identify potential spots for a new German restaurant in New York.

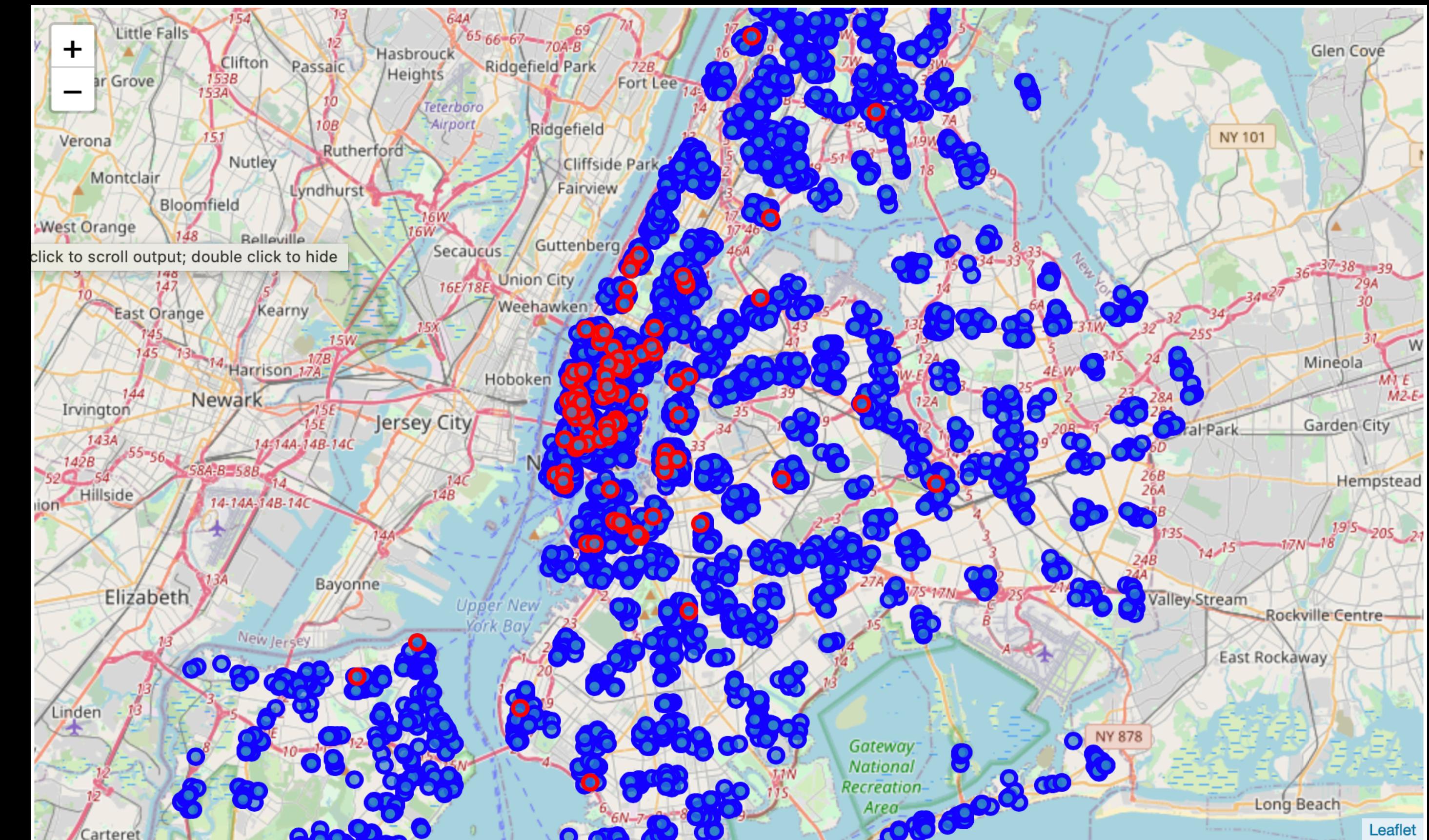
Neighborhoods in New York

- New York has a total of 5 boroughs and 306 neighborhoods.
- In order to identify those neighborhoods in the Foursquare database later, I need a a data set that contains the 5 boroughs and the neighborhoods in them together with the geodata (latitude and logitude coordinates). I download a GEOJSON file with the information from the website https://geo.nyu.edu/catalog/nyu_2451_34572.



Top 100 Food Venues in New York Neighborhoods

A map is good way to get an overview over neighborhoods where a certain amount of restaurants already exist, indicating that there is a customer demand for restaurants in those areas, and where at the same time the number of german restaurants is low or even zero.

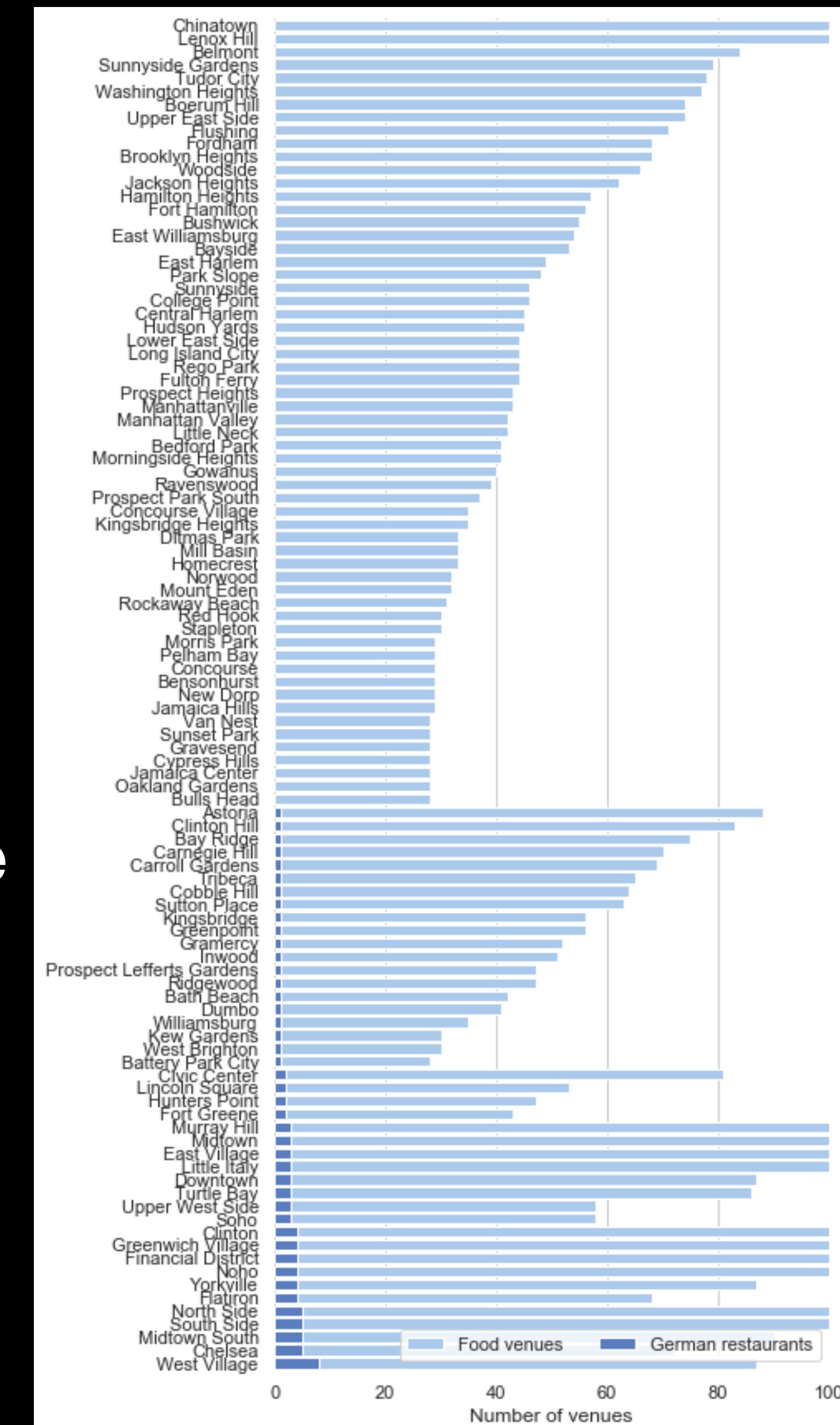


Blue dots indicate a food venue, red dots indicate a German restaurant

Results

German restaurants are still underrepresented in New York

The analysis shows that there are a lot of neighborhoods in New York that have a large number of food venues, indicating that there is a certain demand for food venues and restaurants in these areas while there is no German restaurant - at least no one that Foursquare would recommend based on the user tips.



Discussion of results

- This analysis provides a list of areas with a high density of food venues and low number of German restaurants as a starting point for a further, more customer oriented analysis of the areas.
- While the list and the maps are good starting point to look for areas with high potential for a German restaurant there are certain limitation due to the used dataset. The approach of looking for venues in a certain range around a given geo may lead to incomplete lists of venues. Focussing only on recommended venues also may shorten the list. Nonetheless - if we aim for a restaurant with a great customer experience, low-rated restaurants nearby may be no real competition.
- Further research is needed to analyse the buying power and cultural orientation in the neighborhoods to find the perfect place for a new restaurant. Some of the neighborhoods that have no German restaurant may have none for a reason - from a naive perspective for example China Town may not be the first place to think about opening a German restaurant.

Conclusions

- This test case led to some nice maps and already produced some good results. Real world use cases would probably work with clearer expectations towards the right spot for a new branch and allow to search more goal oriented. Also some of the limitations of the dataset may be corrected by adding additional resources into programming.
- Thank you for reviewing my work. I enjoyed the course but I also clearly see that way more training is necessary to get used to the Python language and work way faster and more efficient.